

BULLETIN OF
THE NEW YORK ACADEMY
OF MEDICINE



OCTOBER 1939

PROTEINS AS CHEMICAL SUBSTANCES
AND AS BIOLOGICAL COMPONENTS

EDWIN J. COHN

Harvey Lecture, January 19, 1939

THE importance of the diverse class of molecules that we know as proteins has in recent years become ever more apparent. Proteins were regarded as chemical substances even a century ago. Thus in 1838, Mulder reported elementary analyses of various proteins, including egg and serum albumins, and estimated minimal molecular weights of over 50,000 on the basis of their sulfur and phosphorus content.¹

The sulfur and phosphorus analyses of Mulder were in error, and so were the atomic weights that he employed. But the concept of proteins as chemical substances is implicit in his calculations.

Three years later† than the number containing Mulder's analyses there appeared a discussion in Liebig's *Annalen* of a communication from the French scientist Denis to Liebig in which is described the separation of blood proteins into those soluble in water, and those which were only soluble in salt solutions, that is, into albumins and globulins.² A century ago, therefore, there was a beginning not only of the characterization of proteins as chemical substances, but the fractionation of tissues into their

† In the same volume is a communication of Dr. Bence-Jones³ on the nitrogen containing foodstuffs of the plant world, on the albumins of the brain, and of the egg yolk, reporting analyses carried out in part with Liebig in the Chemical Laboratory in Giessen.

TABLE I*

	Fibrin		Albumin	
			v. Eiern	v. Serum
Kohlenstoff .	54,56	—	54,48	54,84
Wasserstoff .	6,90	—	7,01	7,09
Stickstoff .	15,72	—	15,70	15,83
Sauerstoff .	22,13	—	22,00	21,23
Phosphor .	0,33	—	0,43	0,33
Schwefel .	0,36	—	0,38	0,68

Auf die kleinere Zahl der Atome von Schwefel und Phosphor berechnet, erhält man:

	Fibrin u. Albumin		Albumin v. Eiern		Albumin v. Serum
	Atome		in 100 Th.		Atome in 100 Th.
Kohlenstoff . . .	400	—	54,90	—	400 — 54,70
Wasserstoff . . .	620	—	6,95	—	620 — 6,92
Stickstoff	100	—	15,89	—	100 — 15,84
Sauerstoff	120	—	21,55	—	120 — 21,47
Phosphor	1	—	0,35	—	1 — 0,35
Schwefel	1	—	0,36	—	2 — 0,72
Atomgewicht =	55692,61	—			= 55893,78.

protein components.

In the intervening century more and more proteins have been isolated from biological systems, purified and characterized by ever improving methods. The reserve proteins of the seeds of plants were presumably the first to be observed in a crystalline state. Until recently the proteins most investigated were the seed globulins, the proteins of the body fluids, such as hemoglobin, the protein of the blood corpuscle which carries oxygen to the tissues, the albumins of serum and of egg white, the casein of milk, the fibrinogen of plasma responsible for the clotting of the blood, the serum globulins, bearers of immunity, and the globulin of muscle, myosin, concerned with muscular contraction. The latter protein has been repeatedly investigated from many points of view and may be said to represent the transition to the study of tissue proteins; that is, to those molecules to which we now look as elements of morphological structure in the body.

That certain hormones and enzymes are protein in nature has only recently become apparent. The iodine-containing hormone of the thy-

* Reproduced from Mulder.¹

roid gland, the active principle of which is thyroxin or some closely related configuration, is thyroglobulin. In thyroglobulin as in hemoglobin, the active configuration is a prosthetic group containing an element—iodine in the one, iron in the other—not found in most proteins.

Another hormone which has been proven to be protein in nature is the insulin effective in the treatment of diabetes, discovered by Banting and crystallized by Abel. The large number of studies in different laboratories on insulin have thus far failed to reveal, as du Vigneaud has pointed out,⁴ any prosthetic group and thus demonstrated that a pure protein may serve as a hormone.

It is over a decade since Sumner⁵ first crystallized an enzyme, the urea splitting urease, and found it to be a protein. Northrop and his colleagues' crystallization of pepsin, pepsinogen, trypsin, trypsinogen, and chymotrypsin,⁶ all digestive enzymes, and Sumner's of catalase⁵ followed, and it has become apparent from much of Northrop's⁶ and of Bergmann's work^{7,8} that the specificity of enzymes depends upon the chemical structure of the protein molecule.

The vitamins were discovered as a consequence of investigations of the nutritional requirements of the body. Therapy consisted in adding vitamins to the diet, but their physiological function and morphological locus remained unknown. Recently it has been demonstrated that riboflavin which is part of the vitamin B₂ complex is the prosthetic group⁹ of the yellow enzyme of Warburg¹⁰ and that the carotenoids, which are known as vitamin A₁ and A₂ are the prosthetic group¹¹ of visual purple, a protein with a molecular weight of about 250,000.¹²

The bearers of immunity have long been associated with the globulin fractions of the serum. In how far the very great specificity of immune reactions inheres in these, in how far in the specific polysaccharides¹³ with which antigens are associated, is still being investigated.

One must now add to the proteins of importance in physiology, pharmacology, morphology and pathology, at least certain of the viruses hitherto the concern of bacteriology. The virus sizes were known as the result of a series of ultrafiltration studies by Zinsser,¹⁴ Bechhold,¹⁵ Elford,¹⁶ and others,¹⁷ but it remained for Stanley^{18,19} to isolate the virus of tobacco mosaic disease, and demonstrate that it was protein in nature. The huge size of tobacco mosaic virus has been deduced from ultracentrifugal studies²⁰ and its long asymmetrical shape has been investigated by the method of double refraction of flow.^{21,22} Like the myosin

of skeletal muscle,²³ tobacco mosaic virus appears to have a length greater than 5,000 Angstroms, a value which may be contrasted with a diameter of 1.5 Angstroms of the carbon atom, and 44 Angstroms for proteins of the molecular weights of pepsin, insulin and lactoglobulin provided they be considered spherical. Not all of the proteins of body fluids are, however, to be regarded as spherical, and not all viruses as elongated. Among the former fibrinogen, the protein involved in the clotting of blood, is elongated and among the latter the newly crystallized Bushy Stunt virus of the tomato²⁴ is nearly spherical.²⁵

The diversity of the functions subserved by protein molecules is not greater than the diversity in their sizes, shapes, and stereochemical configurations. Obviously some of the methods that must be employed for the study of these protein molecules were not envisioned by earlier chemical and biochemical investigators. New methods have been developed as the new molecules were discovered and the numbers of them that have been and are being perfected is increasingly gratifying.

THE POLAR AND NON-POLAR GROUPS OF PROTEINS

Despite the diversity in the form and function of proteins, the number of well recognized chemical configurations which constitute the side chains of the amino acids and therefore of the proteins is by no means large. Without considering here either the methods that are available for the estimation of the amino acid composition of proteins, or the manner in which the peptide chain is ordered in the protein, one may consider the nature of the various configurations which constitute the reactive groups of proteins. In Table II molecular weights are given for a few proteins whose amino acid compositions have been partially apprehended. Only the analytical procedures for sulfur and sulfide sulfur, for cystine and methionine, and for tryptophane and tyrosine have been employed in estimating their minimal molecular weights which have elsewhere been compared^{26,27} with those deduced from the osmotic pressure measurements of Sørensen, Adair and Burk, and the beautiful ultracentrifugal studies of Svedberg and his coworkers.²⁸ The number of sulfhydryl and methionyl, of indole and phenolic hydroxyl groups per mole were then computed. The estimates of imidazol groups derived from histidine analysis and of guanidine groups derived from arginine analysis may be considered fairly accurate. But the estimates of $-\text{NH}_3^+$, equivalent to ϵ -amino groups, may well be low since they are based upon

determinations of lysine, often estimated by difference, and besides do not include the hydroxylysine whose isolation from protein has recently been reported.²⁹

The guanidine group is the most basic of which proteins are possessed, and the properties of the protamines must in large part be considered to depend upon this configuration. The ϵ -amino group dissociates with a strength which varies widely with its position in the molecule, but is in any case also strongly basic, whereas the imidazole group present in such large amounts in hemoglobin dissociates at more nearly neutral reactions. These three configurations alone bear positive charges in acid solutions, are responsible for the cationic properties of the proteins, and are present in such amounts as quantitatively to account for their acid combining capacities.^{30,31}

The dicarboxylic acids yield free carboxyl groups when bound in peptide linkage. However, a portion of the dicarboxylic acids, equivalent to that revealed by determinations of ammonia, are considered to be present in the native protein as amides of the type of asparagine and glutamine. The estimate of carboxyl groups, the only acid groups (if we omit from consideration those of phosphoric acid) which dissociate in relatively acid solution and bear negative charges at neutral reactions must be considered less satisfactory, not only because of the analytical procedures that have thus far been employed, but because the result is a difference between various analytical procedures.

Although only carboxyl groups presumably bear negative charges at acid and neutral reactions, the phenolic hydroxyl group of tyrosine and the sulfhydryl group of cysteine yield negatively charged groups and may therefore be considered to add to the anionic valence of proteins, in so far as these configurations exist^{31,32,33,34,35} in the native molecule. In the case of many proteins there is no evidence of free SH groups in the native state, but in certain others, such as myosin,^{23,33} sulfhydryl groups appear to be free, and these should dissociate at roughly the same alkalinities as free phenolic hydroxyl groups.

The simplest of the amino acids, glycine, contributes neither anionic nor cationic properties to the protein in which it is held in peptide linkage. Nor does it contribute a side chain terminating in a polar or non-polar group. Absence of side chains permits glycine peptides to pack closely in the crystal lattice and presumably is related to their low solubilities. Residues of glycine situated between those of other amino acids

cannot pack equally closely, but rather may be thought of as creating cavities on the surface of the molecule.

Of the side chains which do not lead to ionic configurations the amide groups combined with dicarboxylic acids, the hydroxyl groups of serine, threonine, hydroxyproline, and presumably other hydroxy amino acids, must be considered very polar.* On the other hand, the benzene ring of phenylalanine, and those parts of the pyrrolidine ring of proline that extend from the peptide chain must be considered non-polar.

Aside from its polar nature, the hydroxyl group has significance for the chemistry of the proteins because of its close relation to the solvent water, because of its capacity to enter into ester linkage with acids, being the point of attachment of phosphorus at least in certain phosphoproteins, and because of the possible role, not considered here, of the hydroxyl bond for internal protein structure.

Whereas the hydroxyl groups of aliphatic amino acids and of hydroxyproline, the sulfur of methionine and the nitrogen of the indole group must be considered polar, the configurations of which they are a part reflect also their non-polar groups. Thus tryptophane and methionine behave in many respects like leucine.

The number of CH_2 groups vary in the aliphatic series from glycine to leucine and this will affect interactions with polar and non-polar solvent and solute molecules in various ways.^{36,37} Moreover, the spatial arrangements of amino acid residues of various side chain lengths† must have a determining influence on the distance of closest approach of protein molecules both with ions, dipolar ions, and uncharged molecules.

It is of interest to note the differences in the free groups of different proteins (Table II) and to point out the importance that must be attached to the improvement of analytical procedures upon which our notions of the fine structure of the proteins must ultimately rest. Many

* As judged by the solubility ratio in ethanol and water, $\log N_A/N_0$ (see Table III), asparagine, glutamine, as well as aspartic and glutamic acids, and all the hydroxy acids studied thus far are nearly as polar as glycine, and all more polar than alanine. Thus the values of $\log N_A/N_0$ for glycine, asparagine, glutamine, aspartic and glutamic acids, serine, threonine and hydroxyproline are —3.391, —3.402, —3.466, —2.999, —2.992, —3.362, —3.070, —2.893.

† . . . "the peptides of glycine may be considered as simple cylinders of alternating segments of radii 2.61 Å. in the neighborhood of the CH_2 group, and 2.12 Å. in the neighborhood of the amide group. The average radius on this basis would be 2.32 Å., which is one-half the distance of nearest approach of parallel polypeptide chains."

"Substitution of an alanyl for a glycyl residue in the peptide chain may be considered as adding a branch extending 1.26 Å. from the cylinder. Valine would extend the branch to 2.52 Å., norvaline or leucine to 3.78 Å., and norleucine to 5.04 Å."³⁸ Astbury gives "a 'side-spacing' of the order of $4\frac{1}{2}$ to 5 Å.: in most of the x-ray photographs of proteins"³⁸ (p. 199).

"In terms of the dimension of the CH_2 and CONH groups the packing of peptide chains in the two planes should be given by 4.64 (i.e., 2×2.32) and 9.68 (i.e., 2×4.84) Å., the amino acids of the protein having hydrocarbon chains, on the average, of the length of valine. The presence of amino acids with longer hydrocarbon chains would, of course, still further increase the distance in this plane between parallel polypeptide chains"³⁸ (p. 95).

TABLE II
NATURE AND NUMBER OF POLAR AND NON-POLAR
SIDE CHAIN GROUPS OF CERTAIN PROTEINS

SUBSTANCE	MOLECULAR WEIGHT	CATIONIC GROUPS			ANIONIC GROUPS			NON-IONIC POLAR GROUPS				NON-POLAR GROUPS			
		Guanidine	Ammonium	Imidazole	Carboxyl	Phenolic-hydroxyl	Sulphydryl or $\frac{1}{2}$ (S-S)	Amide	Hydroxyl	Methionyl	Indole	Paraffin	Benzene	Pyrrolidine	Glycine
		Number of Groups													
Glycine	75		1		1										
Cystine	240		2		2		2								
Salmin	5,600	24			1				4			2		6	
Egg Albumin	35,000	12	12	4	28	8	6	27		12	2	45	12	12	
Insulin	35,000	6	6	18	36	24	36	36				81	+	+	
Zein	39,000	4	0	2	10	12	3	84	6	6		122	18	30	
Gliadin	42,000	9	2	9	18	8	8	128	1	6	2	43	6	48	0
Hemoglobin	66,700	16	38	33		12	4		8	4	4	180	18	12	
Serum albumin	73,000	20	66	16	*	18	36	56	4	2	134	14	6		
Casein	96,000	21	42	16	146	+	3	90	6	21	8	156	24	72	6
Edestin	309,000	288	72	42	228	78	36	414	58	48	24	616	60	112	

* The values for the sum of the dicarboxylic acids thus far available are even smaller than demanded by the ammonia analysis. The high base combining capacity of serum albumin suggests the desirability of new studies of the dicarboxylic acids of the serum albumins.

† Too discrepant results to be considered.

In constructing this table the following assumptions have been made relating the groups to the amino acids found on hydrolysis:

Guanidine	= Arginine.
Ammonium	= Lysine.
Imidazole	= Histidine.
Carboxyl	= Aspartic + glutamic + β -hydroxyglutamic acid — ammonia.
Phenolic-hydroxyl	= Tyrosine.
Sulphydryl or $\frac{1}{2}$ (S-S)	= 2 \times Cystine, or alkali-labile sulfur if in excess of cysteine sulfur.
Amide	= Ammonia.
Hydroxyl	= Serine + threonine + hydroxyproline. The numbers of the reported residues of hydroxyglutamic acid are 3 for egg albumin, 6 for zein, 20 for gliadin and 64 for casein, and not included here, but have been counted as carboxyl groups
Methionyl	= Methionine, or total sulfur — alkali-labile sulfur.
Indole	= Tryptophane.
Paraffin	= Alanine + valine + leucine + norleucine. Threonine and methionine are not included since certain of their non-polar groups are situated between polar groups. The total number of CH_2 groups in paraffin side chains (alanine + serine) + 2 threonine + 3 (valine + methionine) + 4 (leucine + norleucine) is for present estimates on these proteins 181, 324, 378, 152, 648, 474, 576, 2246 respectively for egg albumin, insulin, zein, gliadin, hemoglobin, serum albumin, casein and edestin.
Benzene	= Phenylalanine. In tyrosine the benzene rings are situated between polar groups. The number of such benzene rings is given by the number of phenolic hydroxyl groups.
Pyrrolidine	= Proline. In hydroxyproline the non-polar groups are situated between the hydroxyl groups and the peptide chain. The numbers of these, included in the number of hydroxyl groups, are 2 for zein, 4 for hemoglobin, 2 for casein and 48 for edestin. The hydroxyl minus the hydroxyproline groups yield, of course, serine plus threonine.

proteins appear to have molecular weights ranging from 34,000 to 42,000. The analyses on the basis of which their percentile compositions have been determined have been,^{8,26,40} and will be, discussed elsewhere. Although it must be remembered that in many cases the number of groups of any kind may at this time only be estimates, since the analytical evidence regarding many of these residues is by no means satisfactory, there is none the less an advantage in considering the differences in protein molecules which emerge from the analyses at present available.

Zein and gliadin among those of roughly the same molecular weight are classified as prolamines. That is to say, they are insoluble in water, but soluble in alcohol-water mixtures. Their compositions are known to over 90 per cent, and so it would appear to be significant that they contain but very few basic amino acids and very few free carboxyl groups, but consist predominantly of glutamine and asparagine, of proline and of mono-amino monocarboxylic acids with paraffin side chains. Even certain differences between zein and gliadin may be worth noting. Thus the former is poorer in amide and pyrrolidine groups, whereas the latter is poorer in the number of non-polar CH₂ groups in paraffin side chains.

As compared with the six cationic and eight or ten carboxylic groups of zein, and the eighteen cationic and eighteen free carboxylic groups of gliadin, egg albumin has approximately twenty-eight cationic and an equal number of free carboxylic groups and is extremely soluble in water, being, therefore, classified as an albumin. The protein of roughly the same molecular weight, crystallized some years since from the albumin fraction of milk by Palmer,⁴¹ has a solubility of approximately 0.8 gram per liter in water at 25° C., but is relatively soluble in salt solutions and has therefore been called lactoglobulin. Although we already know a great deal about this protein as a result of physical-chemical studies,⁴² analytical studies regarding its amino acid composition remain inadequate.

Two of the other proteins of this size have well known physiological functions: pepsin is an enzyme, insulin a hormone. Pepsin is very poor in basic amino acids, rich in free carboxyl groups and has a very acid isoelectric point. It is also richer in tyrosine than the other proteins thus far considered, though not richer than insulin. Perhaps the outstanding characteristic of the latter protein, however, is the large number of cystine residues that it contains rather than its electrically charged groups, which would not appear on the basis of existing analyses to be

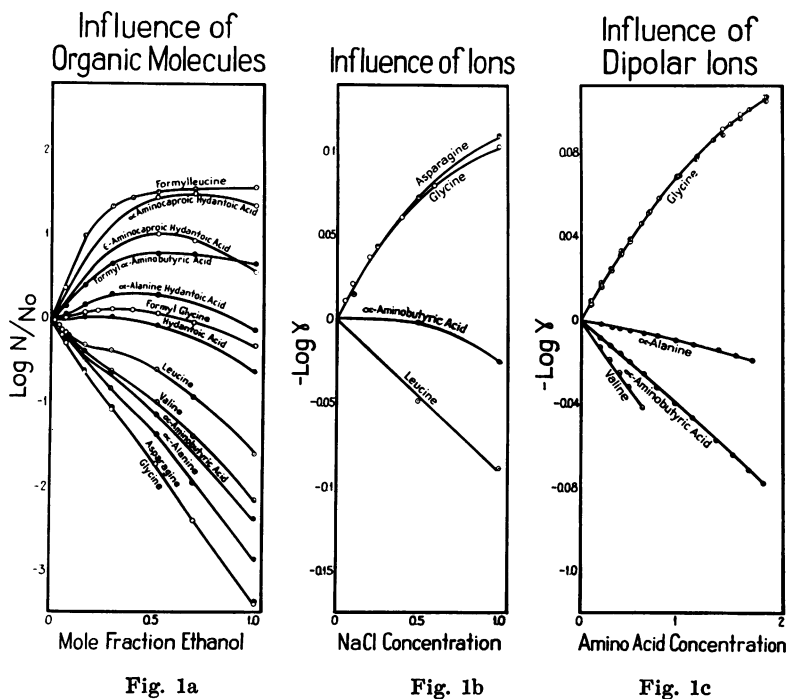
very different from those of other proteins. Pepsin may be classified as a globulin, since it is soluble in salt solutions but not in distilled water. Certainly there are differences in the composition of these proteins of nearly the same size and shape, but certainly also these do not yet suffice to allow us to predict behavior—especially physiological behavior.

The evidence thus far available suggests that the molecular weight in dilute aqueous salt solution of hemoglobin, of casein and of edestin are of the sizes indicated. These molecular weights are many times the minimal molecular weights calculated on the basis of composition, and it has been reported that the molecular weights of these proteins in concentrated urea solutions and also under certain other conditions are far smaller. The number of reactive groups of each kind in these proteins may most readily be compared with those of molecular weights from 34,000 to 42,000 by dividing the results for hemoglobin by two, for casein by three, and for edestin by from six to eight. Among other differences it should be noted that edestin is far richer in guanidine, and hemoglobin in the imidazole group. The imidazole group of the latter has been demonstrated to be in such a relation to the prosthetic iron containing group that its state influences the combination of the protein with oxygen.^{43,44,45}

Physiological behavior may often depend upon a very small fragment of the molecule, the groups of which are arranged in a special configuration. In the case of hemoglobin and of cytochrome, of the hemocyanins, of the yellow enzymes, and of certain other proteins the reactive group may be considered as prosthetic, that is, as not being derived from amino acids. In the case of insulin, however, as we have seen, investigation has thus far revealed no components of the molecule other than the normally occurring amino acids, and the specificity of physiological behavior of this molecule,⁴ as of many enzymes,^{6,7} would thus appear to depend merely upon the arrangement in space of charged and uncharged, polar and non-polar side chains.

INFLUENCE OF POLAR AND NON-POLAR GROUPS UPON SOLUBILITY IN ORGANIC SOLVENTS

Certain correlations can be made between proteins that are rich in cationic groups and in acid combining capacity, between those that are rich in anionic groups and in base combining capacity, and that are poor in ionic groups, and therefore in amphoteric properties. Those among



the latter rich in pyrrolidine and paraffin groups appear also to be more soluble in ethanol-water mixtures than in aqueous solutions and this has been the method for their isolation and the basis for their classification. It has been possible to develop a firm quantitative foundation for these generalizations by studying molecules of known structure containing the various groups of which proteins are possessed.

Glycine is the most soluble in water of the mono-amino mono-carboxylic, aliphatic α -amino acids. The additional CH_2 group of alanine renders it less soluble in water. Valine with 3 CH_2 groups in its side chain is still less soluble,⁴⁶ and the solubility of leucine with 4 such CH_2 groups is 0.0744 mole per liter as contrasted with 2.886 moles per liter for glycine at 25° C. These amino acids are all far less soluble in alcohol-water mixtures than in water, but the effect of alcohol in diminishing solubility is smaller the larger the number of CH_2 groups of which the α -amino acid is possessed. This is illustrated in Fig. 1a in which the logarithm of the solubility in any given solvent, N , divided by that in water, N_0 , is plotted as ordinate and the mole fraction ethanol in the system as abscissa. That the smaller effect of ethanol in diminishing solu-

bility the longer the paraffin side chain depends upon the dipolar ionic nature of the neutral amino acids has been demonstrated by their comparison with certain of their derivatives that have been purified or synthesized by my colleague, T. L. Meekin.

Curves describing the behavior in ethanol-water mixtures of the formyl derivatives of the amino acids and their hydantoic acids,⁴⁷ also form a family, though their shape is quite different from that of the amino acids from which they are derived. Whereas small amounts of alcohol diminish the solubility of α -amino acids, they increase the solubility of their derivatives which are no longer dipolar ions. This is particularly marked with formylleucine and α -aminocaproic hydantoic acid. The additional CH_2 groups of these molecules as compared with hydantoic acid are reflected by increased solubility in systems rich in ethanol. The isomeric α -aminocaproic hydantoic acid, in which the CH_2 groups lie between polar groups behaves far more like the formyl derivative of α -aminobutyric acid⁴⁷; that is to say, like a molecule with two less CH_2 groups. Its solubility in 80 per cent ethanol is approximately tenfold that in water and threefold that in ethanol, a type of behavior characteristic of the prolamines, zein and gliadin.

The influence of each CH_2 group in the transfer from water to ethanol would, however, appear to be the same for the α -amino acids or for their derivatives. This has previously been demonstrated⁴⁶ on the basis of data such as that in Table III by subtracting $\log N_A/N_0$ (where N_A is the solubility in mole fraction in ethanol) of glycine from alanine and from α -aminobutyric acid. The differences per CH_2 group are the same, not only for this comparison but for that of the other amino acids of the same family that have been adequately studied in this way, as well as for derivatives of the amino acids and for isomers of the amino acids, such as the α hydroxyamides, glycolamide and lactamide, which have the same composition though different structures than the amino acids. Moreover, expressed in these terms the effect is always positive and may tentatively be put equal to 0.5. Differently stated *each CH_2 group in side chains terminating in a methyl group may be thought of as increasing solubility in ethanol relative to that in water threefold for each CH_2 group.*

The non-polar CH_2 group does not have this effect when placed between polar groups.⁴⁷ In the peptide chain CH_2 groups alternate with CONH groups and the element CH_2CONH may therefore be con-

sidered as the repeating pattern in the protein and as having a length in the stretched condition of 3.5 Å. as revealed by x-ray studies of the fibrous proteins.^{39,48}

The influence of the configuration CH_2CONH is opposite in sign to that of the CH_2 group as is demonstrated by the studies of three series of compounds in Table III. This configuration, which constitutes the backbone of the protein molecule, as well as certain protein side chains, must therefore be considered strongly polar in nature.⁴⁹

The hydroxyl group must also be considered polar; indeed it is because of this group that the glycerols, alcohols, and water are to be regarded as polar solvents; water being the most polar since it has no CH_2 group.

The influence of the hydroxyl group in proteins on the transfer from water to ethanol is estimated in Table III from a comparison of amides and hydroxyamides, of serine and alanine and of threonine and α -aminobutyric acid.⁴⁹ The comparisons of hydroxyproline and proline and of tyrosine and phenylalanine are omitted, since they include the effect on non-polar carbon rings of their position between polar groups.

The sulfur of methionine is between non-polar groups, and comparison of this molecule with one of the same composition save for the sulfur suggests that the latter diminishes the influence of the paraffin side chain by an amount approximately equal to one CH_2 group. Comparison of phenylalanine with alanine however demonstrates that *the effect of the non-polar benzene ring is of the same sign as the effect of the CH_2 group, but opposite in sign to the CH_2CONH and OH groups.*

Far greater than any of these is the influence of dipolar ionic structure. The demonstration in 1923 by Bjerrum⁵⁰ that amino acids were to be regarded as zwitterions or dipolar ions rendered it possible to understand the very strongly polar nature not only of amino acids, peptides and proteins, but their far higher solubility in water than in non-polar solvents. The amino acids are compared with certain of their derivatives in Fig. 1 and in Table III. The change in $\log N_A/N_o$ may be taken as close to -2.7 , whether amino acids and peptides are compared with hydantoic acids or with α -hydroxy-amides. *The influence of dipolar ionic structure upon this transfer is therefore at least four times as great as that of an hydroxyl group or of the CH_2CONH configuration.* Although the backbone of the protein may thus be conceived of as consisting of alternating polar and non-polar groups, and although many of

TABLE III

INFLUENCE OF STRUCTURE ON SOLUBILITY IN WATER AND ETHANOL AT 25°

<i>Substance</i>	LOG N_A/N_0	LOG N_A/N_0	$\left(\frac{\Delta \log N_A/N_0}{n} \right)$
Influence of CH ₂ group			
Glycine-alanine	-3.391	-2.856	+0.54
Glycine- α -aminobutyric acid	-3.391	-2.375	+0.50
Glycine- α -aminocaproic acid	-3.391	-1.414	+0.49
Valine-leucine	-2.158	-1.622	+0.54
Hydantoic- α -alanine hydantoic acid	-0.630	-0.137	+0.49
Hydantoic- α -aminocaproic hydantoic acid	-0.630	+1.352	+0.49
Glycolamide-lactamide	-0.799	-0.254	+0.54
Glycolamide- α -hydroxycaproamide	-0.799	+1.084	+0.47
Formylglycine-formylaminobutyric acid	-0.330	+0.651	+0.49
Formylglycine-formylleucine	-0.330	+1.556	+0.47
Influence of CH ₂ CONH group			
Glycine-diglycine	-3.391	-4.367	-0.98
Glycine-triglycine	-3.391	-4.965	-0.79
Hydantoic acid-diglycine-hydantoic acid	-0.630	-1.533	-0.90
Hydantoic acid-triglycine-hydantoic acid	-0.630	-2.253	-0.81
Glycolamide-glycolylglycine amide	-0.799	-1.517	-0.72
Influence of OH group			
Acetamide-glycolamide	-0.120	-0.799	-0.68
Propionamide-lactamide	+0.016	-0.254	-0.27
Caproamide- α -hydroxycaproamide	+1.726	+1.084	-0.64
Alanine-serine	-2.856	-3.362	-0.51
α -Aminobutyric acid-threonine	-2.375	-3.070	-0.69
Influence of methionyl sulfur			
α -Aminovaleric acid-methionine	1.90*	-2.444	-0.54
Influence of benzene ring			
Alanine-phenylalanine	-2.856	-1.453	+1.40
Influence of dipolar ionization			
Glycine-glycolamide	-3.391	-0.799	-2.59
Alanine-lactamide	-2.856	-0.254	-2.60
Norleucine- α -hydroxycaproamide	-1.414	+1.084	-2.50
Glycine-hydantoic acid	-3.391	-0.630	-2.76
Alanine-alanine hydantoic acid	-2.856	-0.137	-2.72
α -Aminocaproic acid- α -aminocaproic hydantoic acid	-1.414	+1.352	-2.77
Diglycine-glycolylglycine amide	-4.367	-1.517	-2.85
Diglycine-diglycine hydantoic acid	-4.367	-1.533	-2.83
Triglycine-triglycine hydantoic acid	-4.965	-2.253	-2.71

* Interpolated from α -aminobutyric and α -aminocaproic acids.

the side chains of proteins are polar, the ionic groups and the dipolar ionic structure of proteins must be considered for an adequate understanding of protein behavior.

The properties of the proteins must from this point of view be conceived of as depending upon the properties, the number and the distribution of the polar and the non-polar, the ionic and the uncharged groups of which they are composed, of their volumes, their moments, their arrangements in space, and their reactivities with ions and other dipolar ions.

INFLUENCE OF POLAR AND NON-POLAR GROUPS UPON SOLUBILITY,
AND UPON ACTIVITY COEFFICIENTS OF DIPOLAR
IONS IN SALT SOLUTIONS

The precipitation of proteins by neutral salts has been employed as a method for "their separation, purification, characterization and occasionally classification"⁵¹ ever since the "procedure was first employed in the middle of the last century by Panum,⁵² Virchow⁵³ and Claude Bernard."⁵⁴ Hofmeister demonstrated that "salting-out" depended upon the character of the neutral salt as well as of the protein, and his studies of solubility in concentrated salt solution have since been supplemented by those of Chick and Martin,⁵⁵ Sørensen and his coworkers,⁵⁶ and by various studies from this laboratory.^{51,57,58} Certain of the results in concentrated ammonium sulfate solutions are graphically represented in Fig. 2 in which the logarithm of the solubility is plotted as ordinate and the ionic strength, $T/2$, as abscissa. The linear relation that then obtains was demonstrated in 1925³⁰ and its significance discussed. This "salting-out" relation is characteristic not only of proteins but of gases, of uncharged molecules and, in sufficiently concentrated solutions, of electrolytes. The problem that remained therefore was to determine which groups, or arrangement of groups, of proteins led to their being readily precipitated, which to their being precipitated only from concentrated solutions of certain neutral salts.

On the assumption, implicit throughout this discussion, that amino acids containing the same groups are the prototypes for protein behavior, one may consider the solubility in salt solutions of these smaller molecules of known structure. Pfeiffer and his colleagues^{59,60} demonstrated that the solubility of leucine was decreased with increase in sodium chloride concentration, but that the solubility of glycine and aspartic

acid was increased. Tryptophane and tyrosine⁶¹ behave much like leucine, but the solvent action of sodium chloride upon asparagine is greater than upon glycine, whereas that of α -amino-n-butyric acid is but little influenced by low concentrations but precipitated by high concentrations of sodium chloride.

Certain of these results are graphically represented in Fig. 1b. The solubility in water of leucine is so low that solubility ratios may be considered to yield activity coefficients, but that of glycine is so high that a correction must be made for the activity coefficient of glycine in its own saturated solutions (Fig. 1c). The activity coefficient of glycine in sodium chloride, calculated from the freezing point measurements of Scatchard and Prentiss,⁶² have therefore been substituted in Fig. 1b for the earlier solubility measurements.

All of the amino acids whose activity coefficients are considered in Fig. 1 are α -amino acids and therefore have the same moment due to their dipolar ionic structure. They differ from each other only in the length and nature of their side chains. The "salting-out" effect is greatest for leucine which has the largest number of non-polar CH_2 groups, and begins at low salt concentrations as is the case with such proteins as fibrinogen.⁵⁷

α -Aminobutyric acid with a smaller paraffin side chain is precipitated only from concentrated solutions as is the case with most proteins. That the "salting-out" from aqueous solutions does not depend completely upon the volume of the molecule follows from the comparison of the solubility in salt solution of α -aminobutyric acid with asparagine and glycine. The two former have almost identical apparent molal volumes,⁶³ yet the activity coefficient of asparagine is slightly greater than that of glycine, presumably because of the polar nature of its terminal amide group.

Proteins like amino acids may be precipitated from aqueous solution either by organic solvents, such as alcohol and acetone, or by neutral salts, such as phosphates and sulfates. For the amino acids and presumably for the proteins, precisely the opposite groups are responsible for these effects. The order of precipitation is inverted in Figs. 1a and 1b. Molecules with side chains terminating in non-polar groups, though less soluble in the polar solvent, water, are not so readily precipitated by organic solvents as those of the same dipole moment but whose side chains terminate in polar groups. They are however more readily pre-

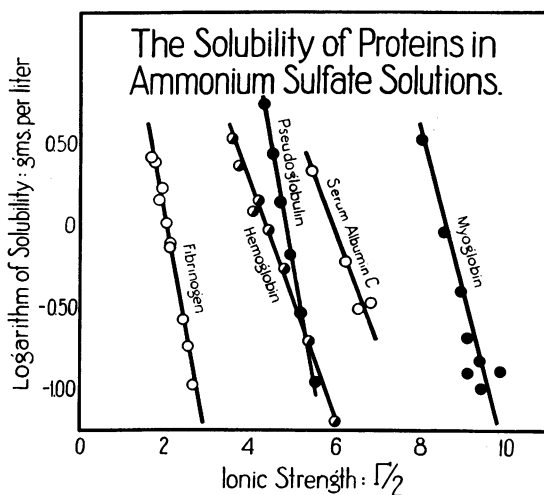


Fig. 2

precipitated by neutral salts. Conversely, molecules constituted predominantly of polar groups are more soluble in the polar solvent, water, from which they are precipitated by organic solvents but dissolved by neutral salts, especially by salts of univalent anions and multivalent cations.

INFLUENCE OF POLAR AND NON-POLAR GROUPS UPON THE ACTIVITY COEFFICIENTS OF DIPOLAR IONS

The interactions of importance in biological systems are not only those between ions and dipolar ions, but also those between the various ionic and especially dipolar ionic species present as components. The present state of experimental and of theoretical knowledge regarding interactions between dipolar ions in aqueous solution have recently been considered elsewhere.³⁷ When one of the dipolar ionic species is but slightly soluble, as in the case of asparagine and cystine, solubility methods have been employed in estimating the activity coefficient of one component in the presence of the other. Such knowledge has now been supplemented by freezing point and vapor pressure measurements upon amino acid solutions, and has yielded the activity coefficients of the single solutes as a function of their concentration.

The simplest dipolar ion, glycine, has been investigated both by freezing point⁶² and by vapor pressure^{64,65} measurements. The results are in excellent agreement and are graphically represented in Fig. 1c

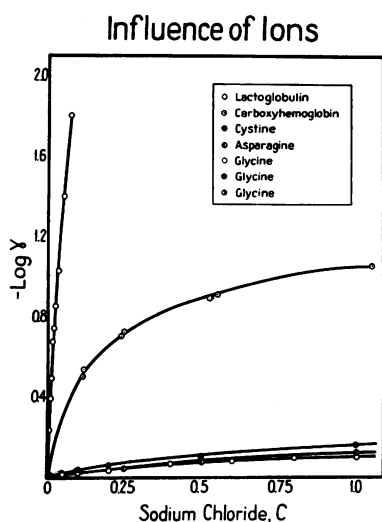


Fig. 3a

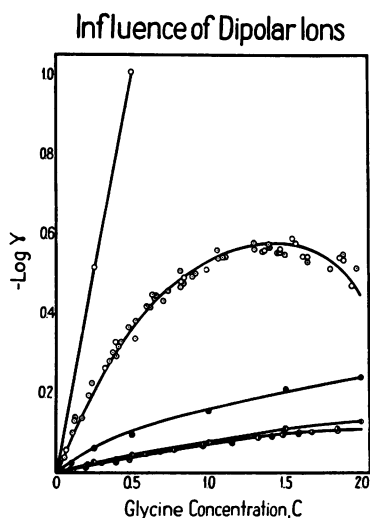


Fig. 3b

in which they are contrasted with activity coefficients calculated from vapor pressure measurements for alanine, α -amino butyric acid and valine. These results may, therefore, be compared with changes in the activity of the same components brought about by neutral salts (Fig. 1b). The comparable nature of the forces between dipolar ions and between ions and dipolar ions is evident. The influence of glycine molecules upon each other, or upon other dipolar ions, is superficially not dissimilar to that of sodium chloride (Fig. 3). *Dipolar ions with paraffin side chains, however, are "salted-out" by each other much as they are by neutral salts.*

In how far the activity coefficient of a dipolar ion will deviate from unity, in how far its interactions with other dipolar ions or with ions will lead to mutual precipitating action or mutual solvent action, will thus depend upon the balance between its non-polar groups and polar groups, especially its dipolar ionic groups.

INFLUENCE OF DIPOLE MOMENT UPON THE ACTIVITY COEFFICIENTS OF DIPOLAR IONS

The polarity of a group is quantitatively measured only in the gaseous state or in non-polar solvents and is generally defined as a displacement of "the center of gravity" of the positive charge from that of the negative charge. The dipole moment is the product of this distance and of

the charge on the electron, 4.8×10^{-10} . The moment of the alcohols due to their hydroxyl group is 1.7×10^{-18} electrostatic units, and that of water is only slightly higher, 1.87×10^{-18} . The esters of the amino acids have been studied in benzene,⁶⁶ and the moments and volumes of the groups that we have been considering have been estimated from these and related studies. The moments of urea, amino acids, peptides and proteins are all far higher than these. Urea has a moment of 5.4×10^{-18} electrostatic units, but those of α -amino acids have been estimated to be approximately 15, of cystine 27, of diglycine 26, triglycine 32, lysyl-glutamic acid 59, of egg albumin 180, of hemoglobin 500, of lactoglobulin 700, and of pseudoglobulin 1300×10^{-18} electrostatic units* on the basis of dielectric constant measurements.

In the preceding discussion only α -amino acids or their derivatives have been considered. The distance of separation between the positively charged ammonium and the negatively charged carboxyl group was therefore always closely the same and equal approximately to 3 Å. This distance of separation is, of course, greater in the peptides, and this is reflected by larger dipole moments. In the case of the proteins the magnitude of the dipole moment depends upon the spatial distribution of the ionic groups. Were the cationic groups all at one edge of the molecule and the carboxylic groups at the other, the dipole moments would be many times larger than those observed. On the other hand, were the electrically charged groups arranged with complete electrical symmetry, the molecule would have no dipole moment. All proteins thus far investigated have dipole moments, and if these are small in comparison with what they might be on the basis of the number of ionic groups, and the dimensions of proteins, they are far larger than those of any other known substances.

The interactions of peptides and neutral salts (in regions of low dielectric constant where the "salting-out" effect is small in comparison with Coulomb forces) have previously been reported and demonstrated to increase with increase in dipole moment.⁷⁰ Moreover, *the logarithm of the activity coefficient was found to increase by somewhat less than the first power of the dipole moment.*

Neutral salts have a profound solvent action upon the class of proteins termed globulins. This effect, first noted by Denis, was investigated

* The electric moments of peptides and proteins have recently been considered in detail;^{31,36,38,67,68} the methods of measurement and of calculation are therefore not considered here. That of lactoglobulin is from measurements by J. D. Ferry, Cohn, Oncley and Blanchard.⁶⁹

in 1905 by W. B. Hardy⁷¹ and by Mellanby⁷² for serum globulin, and by Osborne and Harris⁷³ for edestin. The changes in solubility observed are so large as to form the basis not only for classification but also for methods of purification. Presumably the dipole moments of these molecules are also very large but they have not yet been investigated, nor have preparations of these proteins been available in such a state of purity that their solubility in solution was independent of the amount present in the solid phase.

Hemoglobin and lactoglobulin satisfy this criterion as closely as any proteins thus far investigated. Activity coefficients in aqueous sodium chloride of hemoglobin⁵¹ and lactoglobulin⁴¹ have been calculated from solubility measurements, and are graphically represented in Fig. 3 a, where they are compared with those of glycine, asparagine and cystine. As in the case of the peptides, the logarithm of the activity coefficient increases with the dipole moment, and as a first approximation, by slightly less than the first power of the dipole moment.

These results would lead one to conclude that globulins were proteins of high dipole moment. Their low solubility in water presumably inheres, therefore, in very high crystal lattice energies for this class as for the peptides of glycine.*

The dipole moments of proteins would appear to be one of their most important characteristics, largely defining not only their interactions with neutral salts, but also with other dipolar ions. The solvent action of glycine upon hemoglobin has been investigated by Richards,⁶⁴ and upon lactoglobulin by us.⁶⁹ Activity coefficients calculated from these results are graphically represented in Fig. 3 b in which they are compared with the amino acids, glycine, asparagine and cystine. The similarity in the action of sodium chloride and of glycine is evident, both in the interactions of dipolar ions and of ions and dipolar ions, and $-\log \gamma$ would appear to increase by something less than the first power of the dipole moment.

The presence of greater numbers of paraffin side chains on the dipolar ions decreases these interactions. The interactions between ions and dipolar ions would also be smaller if, as anions, we considered phosphates, sulfates or acetates, or as cations potassium, rubidium or cesium. On the other hand, they would be greater were the anions bromides or

* The peptides of glycine are more insoluble the larger the number of glycine residues in the chain. On the other hand, for molecules such as α and β -alanine and α and ϵ -aminocaproic acids, the isomer of greater dipole moment is the more soluble when there is no difference in density in the solid state.⁴⁶

iodides, or the cations lithium or calcium,* strontium or lanthanum.

INFLUENCE OF DIELECTRIC CONSTANT UPON THE ACTIVITY COEFFICIENTS OF DIPOLAR IONS

The dielectric constant of a solution may be conceived of as a measure of the dipole moments of the polar groups, and the number of these contained per unit of volume. The dielectric constant of a vacuum is thus unity, and that of the alcohol series is larger the smaller the number of non-polar CH_2 groups to which the polar hydroxyl group is attached. Indeed it can be shown that the product of the molal volume and the dielectric constant of solutions of the alcohols from heptanol to methanol is essentially constant^{38,76} and that this relation extends also to the first member of the series, water, which has a dielectric constant of 78.54 at 25° C. as compared with 32.71 for methanol, 24.28 for ethanol and 20.83 for acetone.

But few solvents have higher dielectric constants than water. Among them are hydrocyanic acid and formamide. The latter substance is the first member of the aliphatic amide series, and is in the same relation to acetamide and propionamide as water is to methanol and ethanol. Two of the most polar solvents known, therefore, possess this property by virtue respectively of the hydroxyl and amide groups, both of which, as we have seen, are present in large numbers as constituent parts of protein molecules.

Aqueous solutions of urea have higher dielectric constants than water, and the moment of urea has been estimated, as we have seen, to be 5.4×10^{-18} electrostatic units, or roughly, three times that of the hydroxyl group. The moment of glycine, 15×10^{-18} electrostatic units is, however, three times that of urea, cannot be ascribed to the moments of its polar groups, and is thus evidence of the dipolar ionic structure of the amino acids. The dielectric constant of solutions containing glycine is greater by 23 per mole of glycine per liter of solution. This dielectric constant increment would appear within the limits of measurement to be the same for all α -amino acids, to be greater by 13.3 for each CH_2 group separating the positively charged ammonium from the negatively charged carboxyl group, and by 46.0 for each CH_2CONH group separating the dipolar ionic groups in the series of monoamino-monocarboxylic peptides that have thus far been investigated.^{38,77}

* That is assuming that calcium did not lead to a change in the molecular weight of the protein as appears to be the case for casein⁷⁴ and certain serum globulin fractions.⁷⁵

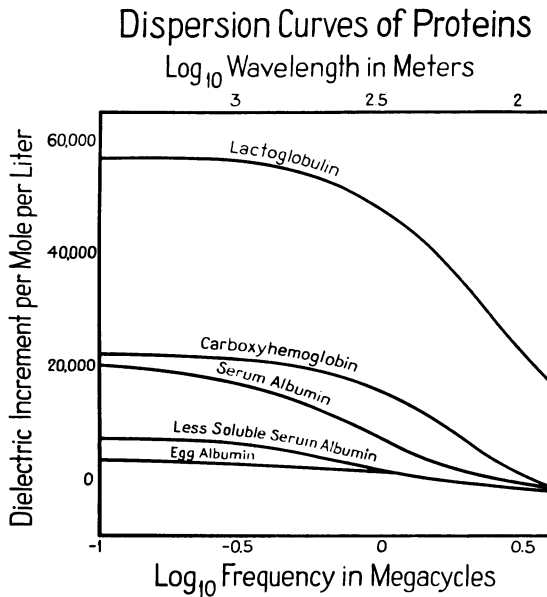


Fig. 4

Peptides have been studied which like cystine contain two positively and two negatively charged groups. The dipole moments of such molecules depend upon the spatial arrangement of the charged groups and so do their dielectric constant increments. This is equally true of the multipolar proteins whose moments have been estimated by means of dielectric constant measurements.

Methods of measuring the dielectric constants of protein solutions have elsewhere been discussed,^{67,78,79} but the significant results of such measurements may well be considered. Placed in an alternating field the isoelectric protein molecules are oriented by virtue of their dipole moments and the capacity of the system and dielectric constant of the solution is thus increased. The curves in Fig. 4 represent this effect, the dielectric constant increments being given by the measurements at low frequency as 3,100 for egg albumin, 22,000 for hemoglobin and 56,000 for lactoglobulin. Of the fractions of crystallized serum albumin that have been investigated⁸⁰ the more soluble has a dielectric constant increment of nearly three times the least soluble fraction thus far investigated.

At higher frequencies the large protein molecules can no longer follow the alternating potential and dispersion of the dielectric constant

results. Plotting the measured dielectric constant against the frequency yields dispersion curves of the kind represented in Fig. 4. At sufficiently high frequencies, the dielectric constant measured is actually smaller than that of the pure solvent, and this effect will be greater the larger the volume of the protein which is no longer able to follow the current and the larger, therefore, the number of the smaller solvent molecules which would be able to follow the current at these frequencies, but which are displaced by the protein molecule.

The critical frequencies approximately represented by the midpoints of the dispersion curves are thus characteristic of the size and shape of the protein molecule, and are functionally related to the relaxation times, a quantity which presumably is of significance in electro-physiology. If proteins have different moments but the same size and shape, they will have the same relaxation time, and this would appear to be the case for all the crystalline serum albumin fractions thus far investigated. Molecules of the same size but different shape will, however, have different relaxation times, and this is true among those whose molecular weight lies between 34,000 and 42,000 of egg albumin and lactoglobulin, and among those whose molecular weight lies between 66,000 and 73,000 of hemoglobin and serum albumin. The latter is presumably the more asymmetric and therefore diffuses more slowly, has a higher viscosity and a higher relaxation time.

Measurements of dielectric constant may thus be employed in estimating not only the dipole moments of proteins, but also their relaxation times, sizes and shapes,^{26,67,80,81} and may be expected to be of great value in the characterization of proteins. Moreover, the dielectric constant of a solution is one of the most important quantities in determining interactions in which electrostatic forces play a role. The high dielectric constants of protein solutions suggest that many systems of biological importance have high dielectric constants, and that an adequate understanding of interactions in certain cells and tissues demands the development of a chemistry for systems of high dielectric constant.

There is a growing body of knowledge regarding the influence upon ions of substances that increase the dielectric constant of aqueous solutions. The influence of amino acids and proteins upon insoluble salts has been explored by the method of solubility^{82,83} and upon soluble salts by the freezing point method,^{62,84} or the electromotive force method.⁸⁵

Rates of reaction depend upon the activity coefficients of the reac-

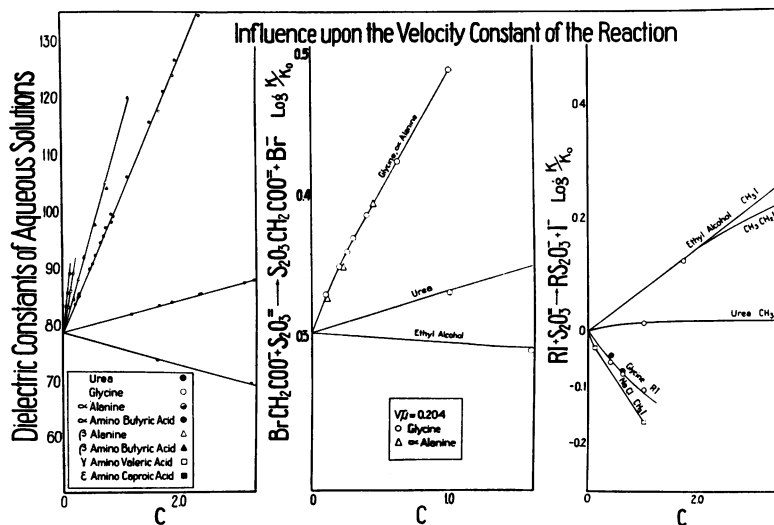


Fig. 5a

Fig. 5b

Fig. 5c

tants. The influence of substances which increase the dielectric constant has been explored upon the reaction of a substance (thiosulfate) on the one hand with uncharged molecules, on the other with ions.⁸⁶ The rates of reaction with the uncharged molecules, increased by ethanol and to a small extent by urea, were found to be decreased by dipolar ions much as they were by ions (Fig. 5c). On the other hand, rates of reaction with ions (Fig. 5b) were found to be increased by urea and dipolar ions as well as by ions. That the influence upon the rate of reaction was not due merely to change in dielectric constant was demonstrated, however, by the study of solutions iso-dielectric with water. The dipolar ions employed were amino acids. Those of greater molecular weight and electrical moment, such as are more characteristic of proteins, remain to be investigated, both from the point of view of their influence upon the environment, and of their role as reactants and as enzymes.

The lowered activity coefficients of ions in dilute aqueous solution have been demonstrated⁸⁷ to depend largely upon electrostatic forces. Such forces, according to Coulomb's law, are inversely proportional to the dielectric constant. In regions of high dielectric constant, therefore, Coulomb forces are diminished and activity coefficients increased. Under these circumstances forces such as those of van der Waals become large in comparison with the greatly diminished Coulomb forces, and differ-

ences in the behavior of ions of the same valence type and of dipolar ions of the same electric moment become apparent.* It is under these circumstances that the so-called Hofmeister series is observed and that great specificity of interaction may be expected.

The influence of dipolar ions upon ions and other dipolar ions in systems of many components will thus be to raise the dielectric constant and thereby to diminish electrostatic forces. Systems containing ions and two dipolar ionic species are the most complex that have thus far been investigated,⁸⁸ but studies upon systems which approach still more closely to those that obtain in nature are in progress, and there would appear to be no differences in principle which should make it impossible to pass from the investigations that have been carried out thus far to the study of interactions between the components of the cells and tissues of the body.

SUMMARY

In order to gain insight into the behavior of even the simplest proteins, we have thus been forced to explore the behavior of molecules of known structure such as the amino acids and peptides. Only those proteins have been considered in this discussion which have been isolated in a relatively pure state, and could therefore be investigated as chemical substances. The methods would appear to be at hand, however, for the study of all proteins. There is no theoretical obstacle to the isolation of all the protein constituents of any given tissue, or to their characterization as chemical substances, and to the study of their interactions as biological components. Proceeding thus, often with new techniques, but employing the classical methods of physical chemistry, we may hope in time to achieve an understanding of the morphology and physiology of biological systems in terms of the properties of their components.

The description of the proteins of tissues as biological components of known size and shape, the number and distribution of their free groups apprehended and the mechanism of their interactions described remain largely to be accomplished. These will depend in many cases upon the spatial configuration of the free groups of the amino acids of

* Conversely, the lower the dielectric constant the greater the Coulomb forces in comparison with the salting-out effect and other non-Coulomb forces. Under these conditions, considered elsewhere,⁹⁰ the principle of the ionic strength obtains in the interactions between ions and dipolar ions, and the valence of the former, and the electrical moments of the latter determine behavior.

which they are composed, and in others upon the presence of more specialized prosthetic groups.

The properties of the proteins as a class permit certain generalizations to be made. By virtue of their large size proteins are retained by cell walls through which oxygen and carbon dioxide, urea and lactic acid, amino acids and many salts freely pass.

By virtue of their ionic nature proteins form salts with each other and with smaller organic and inorganic acids and bases. Certain of these are completely dissociated, others have been judged to be undissociated on the basis of freedom of the non-protein ion to diffuse or to give some chemical test. The precise nature of most of these interactions remains to be investigated.

By virtue of their dipolar ionic structure proteins have high electric moments. These lead, as we have seen, not only to dipole-dipole interactions, but to formation of dipole pairs and to aggregates containing larger numbers of molecules, often arranged with spatial symmetry with respect to each other and dependent upon the shapes of the molecules, the number and distance from the edge of their electrically charged groups and the resulting dipole and multipole moments. As a limiting case the electric moments in the solid state lead to high crystal lattice energies not only in the packing of chemical individuals but of mixed crystals.

By virtue of their high dipole moments proteins contribute high dielectric constants to their solutions. In environments of high dielectric constant Coulomb forces are reduced and conditions obtain that differ widely even from those characteristic of dilute aqueous electrolyte solutions. Although completely dissociated, electrostatic interaction is so reduced that the principle of the ionic strength no longer determines behavior. Under these circumstances, differences in size and shape as well as in valence must be taken into account and non-electrostatic, especially van der Waals forces, acquire primary significance. Under these conditions the behavior of glycine is readily distinguished from that of alanine and of sodium from that of potassium.

By virtue of their high dielectric constants the activity coefficients of both ions and dipolar ions are generally greater than they would be in dilute aqueous solution. As a consequence laws of ideal solution may more often be found to obtain in concentrated biological systems than in dilute aqueous solution. Since rates of reaction are dependent upon

the activity coefficients of the reactants, the high dielectric constant of many biological systems may be expected to play an important role in the kinetics of biological systems.

Various laws in terms of which interionic and intermolecular actions may be described are recent contributions of physical chemistry. Many of them are applicable to biological systems in terms of the parameters that have been discussed. It remains to describe in far greater detail the characteristics of the proteins of biological systems as chemical substances and approach ever more closely to the conditions that obtain in biological systems, and which in large part depend upon the highly specialized physical-chemical characteristics of the protein molecule.

REFERENCES

1. Mulder, G. J. Zusammensetzung von Fibrin, Albumin, Leimzucker, Leucin, u. s. w., *Liebig's Ann. d. Chem. und Pharm.*, 1838, 28: 73.
2. Scherer, J. Chemisch-physiologische Untersuchungen, *Liebig's Ann. d. Chem. und Pharm.*, 1841, 40: 1.
3. Bence-Jones, H. Zusammensetzung der stickstoffhaltigen Nahrungsmittel des Pflanzenreichs, des Albumins und des Eigelbes, *Liebig's Ann. d. Chem. und Pharm.*, 1841, 40: 65.
4. du Vigneaud, V. The chemistry of the hormones from a structural standpoint, *Scient. Monthly*, 1935, 40: 138; and Some aspects of the study of insulin, *J. Washington Acad. Sci.*, 1937, 27: 365.
5. Sumner, J. B. The isolation and crystallization of the enzyme urease, *J. Biol. Chem.*, 1926, 69: 435.
Sumner, J. B. and Dounce, A. L. Crystalline catalase, *ibid.*, 1937, 121: 417.
6. Northrop, J. H. Formation of enzymes, *Physiol. Rev.*, 1937, 17: 144.
7. Bergmann, M. The structure of proteins in relation to biological problems, *Chem. Rev.*, 1938, 22: 423.
8. Bergmann, M. and Niemann, C. On blood fibrin, *J. Biol. Chem.*, 1936, 115: 77; On the structure of proteins, *ibid.*, 1937, 118: 301; and On the structure of silk fibrin, *ibid.*, 1937-38, 122: 577.
9. Theorell, H. Über die Wirkungsgruppe des gelben Ferments, *Biochem. Ztschr.*, 1935, 275: 37; *Enzymforsch.* 1937, 6: 111.
10. Warburg, O. and Christian, W. Über ein neues Oxydationsferment und sein Absorptionsspektrum, *Biochem. Ztschr.*, 1932, 254: 438.
11. Wald, G. Carotenoids and the visual cycle, *J. Gen. Physiol.*, 1935-36, 19: 351; and On rhodopsin in solution, *ibid.*, 1937-38, 21: 795.
12. Hecht, S. and Pickels, E. G. Sedimentation constant of visual purple, *Proc. Nat. Acad. Sci.*, 1938, 24: 172.
13. Heidelberger, M. Chemical nature of immune substances, *Physiol. Rev.*, 1927, 7: 107; and Protein constitution and immunological behavior, *Cold Spring Harbor Symp. Quant. Biol.*, 1938, 6: 369; *Chem. Rev.*, 1939, 24: 323.
14. Zinsser, H. and Tang, F. Studies in ultrafiltration, *J. Exper. Med.*, 1927, 46: 357.
15. Bechhold, J. H. *Die Kolloide in Biologie und Medizin*. Dresden and Leipzig, T. Steinkopf, 4. ed., 1922.
16. Elford, W. J. The sizes of viruses and bacteriophages and methods for their determination: *Handbuch der Virusforschung* (Doerr and Hallauer), Wien, Springer, 1938, p. 126.
17. Ferry, J. D. Ultrafilter membranes and ultrafiltration, *Chem. Rev.*, 1936, 18: 373.
18. Stanley, W. M. Isolation of a crystalline protein possessing the properties of tobacco-mosaic virus, *Science*, 1935, 81: 644.

19. Stanley, W. M. and Loring, H. S. Properties of virus proteins, *Cold Spring Harbor Symp. Quant. Biol.*, 1938, 6: 341.
20. Eriksson-Quensel, I.-B. and Svedberg, T. Sedimentation and electrophoresis of the tobacco-mosaic virus protein, *J. Am. Chem. Soc.*, 1936, 58: 1863.
21. Lauffer, M. A. and Stanley, W. M. Stream double refraction of virus proteins, *J. Biol. Chem.*, 1938, 123: 507; and The viscosity of tobacco virus, protein solutions, *ibid.*, 1938, 126: 443.
22. Mehl, J. W. Double refraction of flow of protein solutions, *Cold Spring Harbor Symp. Quant. Biol.*, 1938, 6: 218.
23. Edsall, J. T. *Personal communication.*
24. Bawden, F. C. and Pirie, N. W. A plant virus preparation in a fully crystalline state, *Nature*, 1938, 141: 513; and Crystalline preparations of tomato bushy stunt virus, *Brit. J. Exper. Path.*, 1938, 19: 251.
25. McFarlane, A. S. and Kekwick, R. A. Physical properties of bushy stunt virus protein, *Biochem. J.*, 1938, 32: 1607.
26. Cohn, E. J. Some physical-chemical characteristics of protein molecules, *Chem. Rev.*, 1939, 24: 203.
27. Cohn, E. J., Hendry, J. L. and Prentiss, A. M. The minimum molecular weights of certain proteins, *J. Biol. Chem.*, 1925, 63: 721.
28. Svedberg, T. Die Molekulargewichts-analyse im Zentrifugalfeld, *Kolloid Ztschr.*, 1934, 67: 2; and Protein molecules, *Chem. Rev.*, 1937, 20: 81.
29. Van Slyke, D. D. *et al.* The unidentified base in gelatin, *Proc. Soc. Exper. Biol. & Med.*, 1938, 38: 548.
30. Cohn, E. J. The physical chemistry of the proteins, *Physiol. Rev.*, 1925, 5: 349.
31. Cohn, E. J. Number and distribution of the electrically charged groups of proteins, *Cold Spring Harbor Symp. Quant. Biol.*, 1938, 6: 8.
32. Mirsky, A. E. and Anson, M. L. Sulfhydryl groups of native proteins—hemoglobin and the proteins of the crystalline lens, *J. Gen. Physiol.*, 1935-36, 19: 439.
33. Greenstein, J. P. Sulfhydryl groups in proteins, *J. Biol. Chem.*, 1938, 125: 501; *ibid.*, 1939, 128: 233.
34. Mirsky, A. E. and Anson, M. L. Sulfhydryl and disulfide groups of proteins, *J. Gen. Physiol.*, 1934-35, 18: 307.
35. Neuberger, A. Electrometric titration of zein and iodozein, *Biochem. J.*, 1934, 28: 1982.
36. Cohn, E. J. Influence of the dielectric constant in biochemical systems, *Chem. Rev.*, 1936, 19: 241.
37. Cohn, E. J., McMeekin, T. L., Ferry, J. D. and Blanchard, M. H. Interactions between dipolar ions in aqueous solution, *J. Physical Chem.*, 1939, 43: 169.
38. Cohn, E. J. The chemistry of proteins and amino acids, *Ann. Rev. Biochem.*, 1935, 4: 93.
39. Astbury, W. T. Some problems in the x-ray analysis of the structure of animal hairs and other protein fibres, *Tr. Faraday Soc.*, 1932, 29: 193.
40. Cohn, E. J. Die physikalische Chemie der Eiweisskörper, *Ergebn. d. Physiol.*, 1931, 33: 781.
41. Palmer, A. H. The preparation of a crystalline globulin from the albumin fraction of cow's milk, *J. Biol. Chem.*, 1934, 104: 359.
42. Cannan, R. K. The effect of neutral salts on the hydrogen ion dissociation curves of protein, *Cold Spring Harbor Symp. Quant. Biol.*, 1938, 6: 1.
43. Cohn, E. J., Green, A. A. and Blanchard, M. H. The amphoteric properties of hemoglobin, *J. Am. Chem. Soc.*, 1937, 59: 509.
44. German, B. and Wyman, J., Jr. The titration curves of oxygenated and reduced hemoglobin, *J. Biol. Chem.*, 1937, 117: 533.
45. Wyman, J., Jr. An analysis of the titration data of oxyhemoglobin of the horse by a thermal method, *J. Biol. Chem.*, 1939, 127: 1.
46. Cohn, E. J., McMeekin, T. L., Edsall, J. T. and Weare, J. H. The solubility of α -amino acids in water and in alcohol-water mixtures, *J. Am. Chem. Soc.*, 1934, 56: 2270.
47. McMeekin, T. L., Cohn, E. J. and Weare, J. H. The solubility of derivatives of the amino acids in alcohol-water mixtures, *J. Am. Chem. Soc.*, 1935, 57: 626; and A

- comparison of the solubility of amino acids, peptides and their derivatives, *ibid.*, 1936, 58:2173.
48. Astbury, W. T. and Woods, H. J. The molecular structure of hair, wool and related fibres, *Nature*, 1931, 127:663; *Proc. Roy. Soc. London*, Ser. B, 1934, 114:314; *Phil. Trans. Roy. Soc. London*, Ser. A, 1933, 232:333.
 49. McMeekin, T. L. *Personal communication*.
 50. Bjerrum, N. Die Konstitution der Ampholyte, *Ztschr. f. physikalische Chem.*, 1923, 104:147; and Dissoziationskonstanten von mehrbasischen Säuren, *ibid.*, 1923, 106:219.
 51. Green, A. A. The solubility of hemoglobin in concentrated salt solutions, *J. Biol. Chem.*, 1931, 93:495; The effect of electrolytes on the solubility of hemoglobin in solutions of varying hydrogen ion activity, *ibid.*, 1931, 93:517; and The solubility of hemoglobin in solution of chlorides and sulfates of varying concentration, *ibid.*, 1932, 95:47.
 52. Panum, P. Neue Beobachtungen über die eisweissartigen Körper, *Virchows Arch. f. path. Anat.*, 1852, 4:419.
 53. Virchow, R. Ueber ein eigenthümliches Verhalten albuminöser Flüssigkeiten bei Zusatz von Salzen, *Virchows Arch. f. path. Anat.*, 1854, 6:572.
 54. Bernard, C., cited by Robin, C. P. and Verdeil, F. *Traité de chimie anatomique et physiologique*, Paris, Baillière, 1853, v. 3, p. 299.
 55. Chick, H. and Martin, C. J. The precipitation of egg-albumin by ammonium sulphate, *Biochem. J.*, 1913, 7:380.
 56. Sørensen, S. P. L. and Horup, M. On the state of equilibrium between crystallized egg-albumin and surrounding mother liquor, *Compt. rend. d. trav. d. Lab. de Carlsberg*, 1915-17, 12:213.
 57. Florkin, M. The solubility of fibrinogen in concentrated salt solutions, *J. Biol. Chem.*, 1930, 87:629.
 58. Morgan, V. F. The solubility of myoglobin in concentrated ammonium sulfate solutions, *J. Biol. Chem.*, 1936, 112:557.
 59. Pfeiffer, P. and Angern, O. Das Aussalzen der Aminosäuren, *Ztschr. physiol. Chem.*, 1924, 133:180.
 60. Pfeiffer, P. and Würzler, T. Die Beeinflussung der Löslichkeit von Aminosäuren durch Neutralsalze, *Ztschr. f. physiol. Chem.*, 1916, 97:128.
 61. von Euler, H. and Rudberg, K. Über Löslichkeitsbeeinflussungen durch Salze, *Ztschr. f. physiol. Chem.*, 1924, 140:113; and Zur Kenntnis der Löslichkeit von Ampholyten, *Ztschr. f. anorgan. u. allg. Chem.*, 1925, 145:58.
 62. Scatchard, G. and Prentiss, S. S. Freezing points of aqueous solutions; mixtures of sodium chloride with glycine and ethyl alcohol, *J. Am. Chem. Soc.*, 1934, 56:2314.
 63. Cohn, E. J., McMeekin, T. L., Edsall, J. T. and Blanchard, M. H. The apparent molal volume and the electrostriction of the solvent, *J. Am. Chem. Soc.*, 1934, 56:784.
 64. Richards, M. M. The effect of glycine upon the activity coefficients of glycine, egg albumin, and carboxyhemoglobin, *J. Biol. Chem.*, 1938, 122:727.
 65. Smith, E. R. B. and Smith, P. K. The activity of glycine in aqueous solution at twenty-five degrees, *J. Biol. Chem.*, 1937, 117:209.
Smith, P. K. and Smith, E. R. B. The activity of aliphatic amino acids in aqueous solution at twenty-five degrees, *ibid.*, 1937, 121:607.
 66. Wyman, J., Jr. and McMeekin, T. L. The dipole moments of esters of amino acids and peptides, *J. Am. Chem. Soc.*, 1933, 55:915.
 67. Williams, J. W. and Watson, C. C. The physical chemistry of the prolamines, *Cold Spring Harbor Symp. Quant. Biol.*, 1938, 6:208.
 68. Wyman, J., Jr. The dielectric constant of solutions of dipolar ions, *Chem. Rev.*, 1936, 19:213.
 69. Ferry, J. D., Cohn, E. J., Oncley, J. L. and Blanchard, M. H. The dielectric properties of lactoglobulin and its interaction with ions and dipolar ions, *J. Biol. Chem.*, 1939, 128:xxvii.

-
70. Cohn, E. J., McMeekin, T. L., Greenstein, J. P. and Wear, J. H. The relation between the activity coefficients of peptides and their dipole moments, *J. Am. Chem. Soc.*, 1936, 58: 2365.
71. Hardy, W. B. Colloidal solution; the globulins, *J. Physiol.*, 1905-06, 33: 251.
72. Mellanby, J. Globulin, *J. Physiol.*, 1905-06, 33: 338.
73. Osborne, T. B. and Harris, I. F. The solubility of globulin in salt solution, *Am. J. Physiol.*, 1905, 14: 151.
74. Philpot, F. J. and Philpot, J. St. L. The effect of calcium on the sedimentation constant of casein, *Proc. Roy. Soc. London*, Ser B, 1939, 127: 21.
75. Svedberg, T. *Personal communication*.
76. Girard, P. Dipole association in pure liquids, *Tr. Faraday Soc.*, 1934, 30: 763.
77. Wyman, J., Jr. and McMeekin, T. L. The dielectric constant of solutions of amino acids and peptides, *J. Am. Chem. Soc.*, 1933, 55: 908.
78. Oncley, J. L., Ferry, J. D. and Shack, J. The measurement of dielectric properties of protein solutions, *Cold Spring Harbor Symp. Quant. Biol.*, 1938, 6: 21.
79. Shutt, W. J. The dielectric capacity of albumin solutions, *Tr. Faraday Soc.*, 1934, 30: 893.
80. Ferry, J. D. and Oncley, J. L. The water-soluble proteins of horse serum, *J. Am. Chem. Soc.*, 1938, 60: 1123.
81. Williams, J. H. and Oncley, J. L. Dielectric constant and particle size, *Physics*, 1932, 3: 314.
82. Failey, C. F. Solubility of thallous chloride in solutions of glycine and glycine salts, *J. Am. Chem. Soc.*, 1932, 54: 576; The solubility of thallous chloride in the presence of edestin nitrate, *ibid.*, 1932, 54: 2367; and Solubility of thallous iodate and thallous chloride in the presence of amino acids, *ibid.*, 1933, 55: 4374.
83. Stone, G. C. H. and Failey, C. F. The activity coefficient of thallous chloride in protein systems, *J. Physical Chem.*, 1933, 37: 935; *J. Biol. Chem.*, 1934, 105: lxxxvi.
84. Kirkwood, J. G. *Personal communication*.
85. Joseph, N. R. The interaction of amino acids and salts, *J. Biol. Chem.*, 1935, 111: 479; 489; and The interactions of calcium chloride and other salts with proteins, as determined by a new type of calcium amalgam electrode, *ibid.*, 1938, 126: 389.
86. Straup-Cope, D. and Cohn, E. J. Influence of amino acids, urea and alcohol upon the velocity constants of chemical reactions, *J. Am. Chem. Soc.*, 1935, 57: 1794.
87. Debye, P. Osmotische Zustandsgleichung und Aktivität verdünnter starker Elektrolyte, *Physikal. Ztschr.*, 1924, 25: 97. Debye, P. and Hückel, E. Gefrierpunkts-erniedrigung und verwandte Erscheinungen, *Physikal. Ztschr.*, 1923, 24: 185.
88. Cohn, E. J., McMeekin, T. L. and Blanchard, M. H. The solubility of cystine in the presence of ions and another dipolar ion, *J. Gen. Physiol.*, 1937-38, 21: 651; *Compt. rend. d. trav. d. Lab. de Carlsberg*, 1938, 22: 142.